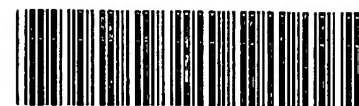O 20

#2

OIPE

ENTERED

RAW SEQUENCE LISTING                    DATE: 03/01/2002
PATENT APPLICATION:  US/10/076,069      TIME: 13:32:40

Input Set : **A:\hepp sequence listing-US.txt**
Output Set: N:\CRF3\03012002\J076069.raw

```
 3 <110> APPLICANT: JURECIC, ROLAND
 4        NACHTMAN, RONALD
 6 <120> TITLE OF INVENTION: HEPP, A NOVEL GENE WITH A ROLE IN HEMATOPOIETIC AND NEURAL
DEVELOPMENT
 8 <130> FILE REFERENCE: 39532-176599
C--> 10 <140> CURRENT APPLICATION NUMBER: US/10/076,069
C--> 10 <141> CURRENT FILING DATE: 2002-02-15
10 <150> PRIOR APPLICATION NUMBER: US 60/268,923
11 <151> PRIOR FILING DATE: 2001-02-16
13 <160> NUMBER OF SEQ ID NOS: 11
15 <170> SOFTWARE: PatentIn version 3.0
17 <210> SEQ ID NO: 1
18 <211> LENGTH: 2082
19 <212> TYPE: DNA
20 <213> ORGANISM: Mus musculus
22 <220> FEATURE:
23 <221> NAME/KEY: CDS
24 <222> LOCATION: (191)...(901)
26 <400> SEQUENCE: 1
27 ccccgcgtcg gtcttccacc tcacctttcg agctggccgc cgcttgctgt gcgcagtttc    60
29 gggggactgg accttccctg gcttttagca gcgccgagcg ccatggcgac cctttgctgg   120
31 gcaggtgacc gattccgggt gcccgaagga gctggcgtgg gtctgccttg cagccgcccg   180
33 cctggacagg atg ttt gct aga ggg ctg aag agg aaa tat ggt gac cag       229
34           Met Phe Ala Arg Gly Leu Lys Arg Lys Tyr Gly Asp Gln
35             1               5                   10
37 gaa gaa gga gta gag ggt ttt ggc act gtc cct tcc tat agc ctg cag     277
38 Glu Glu Gly Val Glu Gly Phe Gly Thr Val Pro Ser Tyr Ser Leu Gln
39     15                  20                  25
41 cga cag tca ctc ctg gac atg tcc ctt gtc aag ctc cag ctc tgt cac     325
42 Arg Gln Ser Leu Leu Asp Met Ser Leu Val Lys Leu Gln Leu Cys His
43 30                  35                  40                  45
45 atg cta gtg gag ccc aat ctc tgc cgc tcg gtc ctc atc gcc aac aca     373
46 Met Leu Val Glu Pro Asn Leu Cys Arg Ser Val Leu Ile Ala Asn Thr
47             50                  55                  60
49 gtc cgg cag atc cag gag gaa atg agc cag gat ggt gtg tgg cat ggg     421
50 Val Arg Gln Ile Gln Glu Glu Met Ser Gln Asp Gly Val Trp His Gly
51             65                  70                  75
53 atg gca ccc cag aat gta gat cgg gca cca gtt gaa cgc ctg gtg tcc     469
54 Met Ala Pro Gln Asn Val Asp Arg Ala Pro Val Glu Arg Leu Val Ser
55          80                  85                  90
57 aca gag atc ctg tgt cgt aca gtg agg gga gct gag gaa gag cac cct     517
58 Thr Glu Ile Leu Cys Arg Thr Val Arg Gly Ala Glu Glu Glu His Pro
59    95                  100                 105
61 gct cct gaa ctg gaa gat gct ccc ttg caa aac tcg gtt tcc gag ctc     565
```

RAW SEQUENCE LISTING                      DATE: 03/01/2002
PATENT APPLICATION:   US/10/076,069       TIME: 13:32:40

Input Set : A:\hepp sequence listing-US.txt
Output Set: N:\CRF3\03012002\J076069.raw

```
62 Ala Pro Glu Leu Glu Asp Ala Pro Leu Gln Asn Ser Val Ser Glu Leu
63 110               115               120               125
65 ccc atc gtt ggc tca gca cca ggg caa agg aac cct cag agc agc ctc   613
66 Pro Ile Val Gly Ser Ala Pro Gly Gln Arg Asn Pro Gln Ser Ser Leu
67               130               135               140
69 tgg gag atg gac agc cca caa gaa aac agg gga agc ttt cag aag tca   661
70 Trp Glu Met Asp Ser Pro Gln Glu Asn Arg Gly Ser Phe Gln Lys Ser
71               145               150               155
73 ctg gac cag ata ttt gag acc ctg gag aac aaa aac tcc agt tca gtg   709
74 Leu Asp Gln Ile Phe Glu Thr Leu Glu Asn Lys Asn Ser Ser Ser Val
75           160               165               170
77 gag gaa ctc ttc tca gat gtg gac agc tcc tac tat gac ctg gac aca   757
78 Glu Glu Leu Phe Ser Asp Val Asp Ser Ser Tyr Tyr Asp Leu Asp Thr
79       175               180               185
81 gtg cta aca gga atg atg agt ggg acc aag tcc agt ctc tgc aat ggc   805
82 Val Leu Thr Gly Met Met Ser Gly Thr Lys Ser Ser Leu Cys Asn Gly
83 190               195               200               205
85 ctt gag ggc ttt gct gca gcc acc cct cct ccc agt tcc act tgc aag   853
86 Leu Glu Gly Phe Ala Ala Ala Thr Pro Pro Pro Ser Ser Thr Cys Lys
87               210               215               220
89 tct gac ctg gct gag ctg gac cat gtg gta gag att ctg gtg gag acc   901
90 Ser Asp Leu Ala Glu Leu Asp His Val Val Glu Ile Leu Val Glu Thr
91               225               230               235
93 tgagaggcca ccccagtggg ctaagggtga ggccaccagt ccccatggag ctcacgtgtg  961
95 ttgtgaccca gagacagata agcacttgtc ctaagagggg ctctggctct tgagctcatt 1021
97 atcctttttgt gtgacattgg actcactgtg gaggatggtg tgtcacagct atgtctagtc 1081
99 tattttcaat tagataggtg aactttctaa aattaagttt tatatgtttt tgggcaatat 1141
101 tttgtcttaa gatatatttt ttaaactttt tatactttag attttttttca gctattttct 1201
103 taaaagtata ttttttctac aaacatcctc tgctgctaca ttagaaacat ttataaccta 1261
105 aatacgattg gtgtgtcatt ttaaaggttt aaatagaaaa cttcttttgt tactgagtct 1321
107 ctacactccc aaggcaactg taaatgtagc cggccgggtg tttacatgag aggctccagt 1381
109 atggtctaca ttctagtaga gcttgaaaag aaccatgcac agctccactg ccccctcact 1441
111 gggtctgctc tggcggatcg gagctctctt cctagccccg tgtgcaggat ggctttattt 1501
113 atgcctattt atatgtaaat gccactgaaa gctaaggtct tactcctgga aatcccaaca 1561
115 ccagttcttc agggactgct gtgaggcagt gccttatgca ggtcttgtcc ttggccatca 1621
117 ctgtctggtt cccagcccag cacatgtgac atgaggacat gacatgcccg aaccacccag 1681
119 caccacatgc tccatgtcaa gtgtgtacgt ggagaccact ggctcccagg cctgtgctca 1741
121 gagagggtgt gcagtcctac gtgtgctggg ggggacgacg gtgacctgtg cttgcttgct 1801
123 tttaaaatgg tgcttggacg ttttaaggtt aaaaacaatc cgactccata tgatttaggg 1861
125 ctcctccacc ctggggtggc ccctatgctg tctgcttgga tctcaaagtc ttggtactcg 1921
127 gcactgtcag actccacccc atgtatcctt tttgtttctc ttgtgctttt tttggacttc 1981
129 ccaacctgag cctaaggttt tatttttatat gtgcttcaat atcaacaatg taaacctcac 2041
131 tttattaaaa gtatccagca aatggaaaaa aaaaaaaaaa a                      2082
134 <210> SEQ ID NO: 2
135 <211> LENGTH: 237
136 <212> TYPE: PRT
137 <213> ORGANISM: Mus musculus
139 <400> SEQUENCE: 2
141 Met Phe Ala Arg Gly Leu Lys Arg Lys Tyr Gly Asp Gln Glu Glu Gly
```

RAW SEQUENCE LISTING                    DATE: 03/01/2002
PATENT APPLICATION:  US/10/076,069      TIME: 13:32:40

Input Set : **A:\hepp sequence listing-US.txt**
Output Set: **N:\CRF3\03012002\J076069.raw**

```
142 1                    5                    10                    15
145 Val Glu Gly Phe Gly Thr Val Pro Ser Tyr Ser Leu Gln Arg Gln Ser
146              20                    25                    30
149 Leu Leu Asp Met Ser Leu Val Lys Leu Gln Leu Cys His Met Leu Val
150          35                    40                    45
153 Glu Pro Asn Leu Cys Arg Ser Val Leu Ile Ala Asn Thr Val Arg Gln
154      50                    55                    60
157 Ile Gln Glu Glu Met Ser Gln Asp Gly Val Trp His Gly Met Ala Pro
158 65                    70                    75                    80
161 Gln Asn Val Asp Arg Ala Pro Val Glu Arg Leu Val Ser Thr Glu Ile
162              85                    90                    95
165 Leu Cys Arg Thr Val Arg Gly Ala Glu Glu Glu His Pro Ala Pro Glu
166          100                   105                   110
169 Leu Glu Asp Ala Pro Leu Gln Asn Ser Val Ser Glu Leu Pro Ile Val
170          115                   120                   125
173 Gly Ser Ala Pro Gly Gln Arg Asn Pro Gln Ser Ser Leu Trp Glu Met
174      130                   135                   140
177 Asp Ser Pro Gln Glu Asn Arg Gly Ser Phe Gln Lys Ser Leu Asp Gln
178 145                   150                   155                   160
181 Ile Phe Glu Thr Leu Glu Asn Lys Asn Ser Ser Ser Val Glu Glu Leu
182              165                   170                   175
185 Phe Ser Asp Val Asp Ser Ser Tyr Tyr Asp Leu Asp Thr Val Leu Thr
186          180                   185                   190
189 Gly Met Met Ser Gly Thr Lys Ser Ser Leu Cys Asn Gly Leu Glu Gly
190          195                   200                   205
193 Phe Ala Ala Ala Thr Pro Pro Pro Ser Ser Thr Cys Lys Ser Asp Leu
194      210                   215                   220
197 Ala Glu Leu Asp His Val Val Glu Ile Leu Val Glu Thr
198 225                   230                   235
```

201 <210> SEQ ID NO: 3
202 <211> LENGTH: 2082
203 <212> TYPE: DNA
204 <213> ORGANISM: Homo sapiens
206 <220> FEATURE:
207 <221> NAME/KEY: CDS
208 <222> LOCATION: (75)..(797)
210 <400> SEQUENCE: 3

```
211 gggaagctgg cggcacagcc gtggcgcctg gctgagcaga ggacccggcg ggcggcctcg      60
213 cgggtcagga caca atg ttt gca cga gga ctg aag agg aaa tgt gtt ggc      110
214                Met Phe Ala Arg Gly Leu Lys Arg Lys Cys Val Gly
215                 1                5                    10
217 cac gag gaa gac gtg gag gga gcc ctg gcc ggc ttg aag aca gtg tcc      158
218 His Glu Glu Asp Val Glu Gly Ala Leu Ala Gly Leu Lys Thr Val Ser
219          15                    20                    25
221 tca tac agc ctg cag cgg cag tcg ctc ctg gac atg tct ctg gtg aag      206
222 Ser Tyr Ser Leu Gln Arg Gln Ser Leu Leu Asp Met Ser Leu Val Lys
223      30                    35                    40
225 ttg cag ctt tgc cac atg ctt gtg gag ccc aac ctg tgc cgc tca gtc      254
226 Leu Gln Leu Cys His Met Leu Val Glu Pro Asn Leu Cys Arg Ser Val
```

```
227 45                    50                   55                   60
229 ctc att gcc aac acg gtc cgg cag atc caa gag gag atg acg cag gat    302
230 Leu Ile Ala Asn Thr Val Arg Gln Ile Gln Glu Glu Met Thr Gln Asp
231                      65                   70                   75
233 ggg acg tgg cgc aca gtg gca ccc cag gct gca gag cgg gcg ccg ctc    350
234 Gly Thr Trp Arg Thr Val Ala Pro Gln Ala Ala Glu Arg Ala Pro Leu
235                      80                   85                   90
237 gac cgc ttg gtc tcc acg gag atc ctg tgc cgt gca gcg tgg ggg caa    398
238 Asp Arg Leu Val Ser Thr Glu Ile Leu Cys Arg Ala Ala Trp Gly Gln
239              95                  100                  105
241 gag ggg gca cat cct gct cct ggc ttg ggg gac ggc cac aca cag ggt    446
242 Glu Gly Ala His Pro Ala Pro Gly Leu Gly Asp Gly His Thr Gln Gly
243       110                  115                  120
245 cca gtt tct gac ctt tgc cca gtc acc tca gca cag gca cca agg cac    494
246 Pro Val Ser Asp Leu Cys Pro Val Thr Ser Ala Gln Ala Pro Arg His
247 125                  130                  135                  140
249 ctg cag agc agc gcc tgg gag atg gat ggc cct cga gaa aac aga gga    542
250 Leu Gln Ser Ser Ala Trp Glu Met Asp Gly Pro Arg Glu Asn Arg Gly
251                      145                  150                  155
253 agc ttt cac aag tca ctt gat cag ata ttt gaa acg ctg gag act aaa    590
254 Ser Phe His Lys Ser Leu Asp Gln Ile Phe Glu Thr Leu Glu Thr Lys
255              160                  165                  170
257 aac ccc agc tgc atg gaa gag ctg ttc tca gac gtg gac agc ccc tac    638
258 Asn Pro Ser Cys Met Glu Glu Leu Phe Ser Asp Val Asp Ser Pro Tyr
259              175                  180                  185
261 tac gac ctg gac aca gta ctg aca ggc atg atg ggg ggt gcc agg ccg    686
262 Tyr Asp Leu Asp Thr Val Leu Thr Gly Met Met Gly Gly Ala Arg Pro
263       190                  195                  200
265 ggc ccc tgc gaa ggg ctc gag ggc ttg gct ccg gcc acc cca ggc cct    734
266 Gly Pro Cys Glu Gly Leu Glu Gly Leu Ala Pro Ala Thr Pro Gly Pro
267 205                  210                  215                  220
269 agc tcc agc tgc aag tcc gac ctg ggc gag ctg gac cac gtg gtg gag    782
270 Ser Ser Ser Cys Lys Ser Asp Leu Gly Glu Leu Asp His Val Val Glu
271                      225                  230                  235
273 atc ctg gtg gag acc tgagcaggag ccctgagtgc tcacagccgc ctctgacgca    837
274 Ile Leu Val Glu Thr
275                      240
277 ttgacacgtg agcactggct cccacggagg gtgcgcctgc cgccagcggc ccagccttgc    897
279 tgccctgtct gctgattctg agaaatccca gaacagccca ttaccagtgg ggctgcagcc    957
281 ctaggcccgt cccactcacc tcccccctgt ggagcgccag gcagaggctg ttctggaagg   1017
283 cttcttgtct tctgacgtcc ccacagccct gggcccctcg tgtctctttg tgtcccccac   1077
285 tgtagaggac ggtgagccgc agctgcatca acctcctttt acctttagat aggtgaattt   1137
287 ttacaattca gttttacatg ttttgggcag tattttgtct taagatatat tttttaaact   1197
289 ttttatacct tatctcttta gattttttca gctattttct taaaagtata tttttttctat  1257
291 aaacatcctt tgctgctaca ttagaacttt tatagcctaa acaattgcag ttggtgtgtt   1317
293 tcattttttt aaggtttaaa taagggtttt ttgtttttgtt ttgttttttg cagtgagcat  1377
295 cactacagtc tcagtcaaca gtgtgaatgt atcatgtttt actttaaatg tgtgtgtgat   1437
297 acttcttcat tatgtcctgc gctgcagtga gacctgggtg aaaatcagga gccgcacaca   1497
299 gccacatctt cctagaccta agagtaaatt atggaggatt ttatttatgt ctatttatat   1557
```

RAW SEQUENCE LISTING          DATE: 03/01/2002
PATENT APPLICATION:  US/10/076,069      TIME: 13:32:40

Input Set : **A:\hepp sequence listing-US.txt**
Output Set: **N:\CRF3\03012002\J076069.raw**

```
301 gtaaatgtca ttgaagacaa aggtcaaata tttgtctgtt tgtagatcac aggcaccagt  1617
303 tggtcttcag ggacctcata gcccctcggt ggtgccttct caaggcagtg ttcctggagg  1677
305 ctcccatcag ggtcagccca tgcacctgcc ctgggtgagg aagtagcatt gctgctggat  1737
307 gagaaacgcc tgcgctgctc tgttagactg gtgctgaaac aaaaggttaa ggctaggttg  1797
309 aagtctagaa tgaaagaaat ctgaatccat gtcattcata accccttgat ctgtagtgtc  1857
311 atgggtgctg ccgcaggcag ggagtgagct ggggtgcct gcagccttcc actcctgccc  1917
313 cgcctcaccc cacatgctcc ctgtttctca tgctttctct aacttcctca cccttaacc  1977
315 aaaaaggtgt gttttctttt gtgcatatag ccattcttaa atatcagtga tgtaaacctc  2037
317 actttattaa aaaattatcc agcaaaaaaa aaaaaaaaaa aaaaa            2082
320 <210> SEQ ID NO: 4
321 <211> LENGTH: 241
322 <212> TYPE: PRT
323 <213> ORGANISM: Homo sapiens
325 <400> SEQUENCE: 4
327 Met Phe Ala Arg Gly Leu Lys Arg Lys Cys Val Gly His Glu Glu Asp
328 1               5                   10                  15
331 Val Glu Gly Ala Leu Ala Gly Leu Lys Thr Val Ser Ser Tyr Ser Leu
332             20                  25                  30
335 Gln Arg Gln Ser Leu Leu Asp Met Ser Leu Val Lys Leu Gln Leu Cys
336         35                  40                  45
339 His Met Leu Val Glu Pro Asn Leu Cys Arg Ser Val Leu Ile Ala Asn
340         50                  55                  60
343 Thr Val Arg Gln Ile Gln Glu Glu Met Thr Gln Asp Gly Thr Trp Arg
344 65                  70                  75                  80
347 Thr Val Ala Pro Gln Ala Ala Glu Arg Ala Pro Leu Asp Arg Leu Val
348             85                  90                  95
351 Ser Thr Glu Ile Leu Cys Arg Ala Ala Trp Gly Gln Glu Gly Ala His
352             100                 105                 110
355 Pro Ala Pro Gly Leu Gly Asp Gly His Thr Gln Gly Pro Val Ser Asp
356         115                 120                 125
359 Leu Cys Pro Val Thr Ser Ala Gln Ala Pro Arg His Leu Gln Ser Ser
360         130                 135                 140
363 Ala Trp Glu Met Asp Gly Pro Arg Glu Asn Arg Gly Ser Phe His Lys
364 145                 150                 155                 160
367 Ser Leu Asp Gln Ile Phe Glu Thr Leu Glu Thr Lys Asn Pro Ser Cys
368             165                 170                 175
371 Met Glu Glu Leu Phe Ser Asp Val Asp Ser Pro Tyr Tyr Asp Leu Asp
372             180                 185                 190
375 Thr Val Leu Thr Gly Met Met Gly Gly Ala Arg Pro Gly Pro Cys Glu
376             195                 200                 205
379 Gly Leu Glu Gly Leu Ala Pro Ala Thr Pro Gly Pro Ser Ser Ser Cys
380         210                 215                 220
383 Lys Ser Asp Leu Gly Glu Leu Asp His Val Val Glu Ile Leu Val Glu
384 225                 230                 235                 240
387 Thr
391 <210> SEQ ID NO: 5
392 <211> LENGTH: 106
393 <212> TYPE: PRT
394 <213> ORGANISM: Danio rerio
```

VERIFICATION SUMMARY                          DATE: 03/01/2002
PATENT APPLICATION:   US/10/076,069           TIME: 13:32:41

            Input Set : **A:\hepp sequence listing-US.txt**
            Output Set: **N:\CRF3\03012002\J076069.raw**

L:10 M:270 C: Current Application Number differs, Replaced Current Application No
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date